Cent. Eur. J. Med. • 8(1) • 2013 • 1-15
DOI: 10.2478/s11536-012-0052-6

VERSITA

Central European Journal of **Medicine**

Prediction of antimicrobial activity of imidazole derivatives by artificial neural networks

Research Article

Małgorzata Wnuk[1], Michał Piotr Marszałł[2], Anna Zapęcka[1], Alicja Nowaczyk[3], Jerzy Krysiński[4], Jerzy Romaszko[5], Piotr Kawczak[6], Tomasz Bączek[6], Adam Buciński[1]*

1 Department of Biopharmacy, Faculty of Pharmacy, Collegium Medicum, Nicolaus Copernicus University, Jurasza 2, 85-094 Bydgoszcz, Poland

2 Department of Medicinal Chemistry, Faculty of Pharmacy, Collegium Medicum, Nicolaus Copernicus University, Jurasza 2, 85-094 Bydgoszcz, Poland

3 Department of Organic Chemistry, Faculty of Pharmacy, Collegium Medicum, Nicolaus Copernicus University, Jurasza 2, 85-094 Bydgoszcz, Poland

4 Department of Pharmaceutical Technology, Faculty of Pharmacy, Collegium Medicum, Nicolaus Copernicus University, Jurasza 2, 85-094 Bydgoszcz, Poland

5 NZOZ Pantamed Sp z o.o. in Olsztyn, 10-461 Olsztyn, Poland

6 Department of Pharmaceutical Chemistry, Faculty of Pharmacy, Medical University of Gdańsk, Hallera 107, 80-416 Gdańsk, Poland

Received 23 February 2012; Accepted 25 April 2012

Abstract: The main goal of our study is the analysis of data obtained from molecular modeling for a series of imidazole derivatives that possess strong antifungal activity. The research was designed to use artificial neural network (ANN) analysis to determine quantitative relationships between the structural parameters and anti-Streptococcus pyogenes activity of a series of imidazole derivatives. ANN in association with quantitative structure-activity relationships (QSAR) represents a promising tool in the search for drug candidates among the practically unlimited number of possible derivatives. In this work, a series of 286 imidazole derivatives presented as cationic three-dimensional structures was used. The activity was expressed as a logarithm of the reciprocal of the minimal inhibitory concentrations, log 1/MIC. Multilayer perceptron ANN was used for predictions of antimicrobial potency of new imidazole derivatives on the basis of their structural descriptors. The obtained correlation coefficient equaled 0.9461 for the learning set, 0.9060 for the validation set and 0.8824 for the testing set of imidazole derivatives. Hence, satisfactory and practically useful predictions of anti-Streptococcus pyogenes activity for a series of imidazole derivatives was obtained, supporting the future successful interpretation of QSAR analysis for those compounds.

Keywords: *Antimicrobial activity • Artificial neural networks (ANN) analysis • Molecular descriptors • Imidazole derivatives*

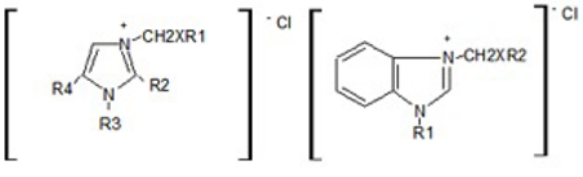
* *E-mail: kizbiofarmacji@cm.umk.pl*

1. Introduction

Current pharmacotherapy still requires novel antimicrobiological medicines in order to effectively fight the risks caused by micro-organisms. An increasing number of strains of microorganisms resistant to known compounds have motivated scientists to undertake further research. In that respect, an interesting group of compounds are azols, particularly imidazoles (Figure 1) and their derivatives, which possess strong antibacterial and antifungal activity. They are important compounds due to their roles in biological systems, particularly when considering enzymes as proton donors or acceptors, coordination system ligands and the base for charge-transfer processes. These molecules belong to the group of surface active compounds, so they have proven antiseptic properties and are used to disinfect sanitary surfaces [1]. The mechanism of action of quaternary ammonium compounds depends on the disruption of the cell membrane and allowance of the connection of a positive-charged group with a negative-charged phosphatic group of phospholipids. In that way, cells of bacteria are isolated from the medium and the multiplication is inhibited. Dependence on structure compounds is related to various types of antimicrobial activity [2-4]. The significance of the position of a substituent in a compound structure for biological activity was proven when investigators evaluated antibacterial activity of two analogs of imidazole derivatives, 5-nitroimidazole and 4-nitroimidazole. A different influence of those analogs on two other groups of micro-organism was also noted [5].

Streptococci are gram-positive, spherical or ovoid cells arranged in chains or pairs. They belong to Lancefield serogroup A, also known as Group A streptococci (GAS). Many species of streptococci are members of the commensial microflora; they tend to colonize the upper respiratory tract and are highly virulent as they overcome the host defense system. *Streptococcus pyogenes* causes diseases such as severe invasive infections, the post-streptococcal sequelae of acute rheumatic fever and rheumatic heart disease, acute glomerulonephritis, and uncomplicated pharyngitis pyoderma, as well as cellulitis, bacteremia, necrotizing fasciitis, and toxic shock syndrome [6-8]. *S. pyogenes* is sensitive to the action of imidazole salts, and those compounds have been specifically analyzed to search for new and efficient drugs [9,10].

Figure 1. The general chemical structure of imidazole derivatives.



a) imidazole ring (R1, R2, R3, R4 – kind of substitute, X – oxygen or sulphur).

b) benzimidazole ring (R1, R2 – kind of substitute, X – oxygen or sulphur).

An artificial neural network (ANN) is a mathematical model analysis that is inspired by the way biological nervous systems, such as the brain, process information. Neural networks are non-linear statistical data modeling tools and can be used to model complex relationships between inputs and outputs or to find patterns in data. An ANN is configured for a specific application, such as pattern recognition or data classification, through a learning process. The increasing use of those models in sciences such as chemistry and biology has been noted since the 1980s. ANNs have been applied in the identification of potential drug targets, modeling of QSAR, compound classification and identification of potential drug candidates [11-15].

Quantitative structure-activity relationship (QSAR) represents an advance by which structural descriptors of a compound are quantitatively correlated with a well defined process, such as biological activity or chemical reactivity. These molecular descriptors are determined empirically and most often by computational methods characterizing physicochemical, pharmacological and toxicological properties of compounds [16,17].

We analysed anti-*Streptococcus pyogenes* activity of minimum inhibitory concentranion (MIC) and three-dimensional QSAR studies for a series of imidazole derivatives. The compounds have been reported in recent literature [2,17] and belong to the class of quaternary ammonium salts used as disinfectants. The aim of our research was to determine the quantitative relationships between structural parameters and the antimicrobial activity of a series of imidazole derivatives with the use of artificial neural networks. In that way, ANN in association with QSAR could represent a promising tool in the search for drug candidates among the practically unlimited number of possible derivatives.

2. Material and methods

2.1. Imidazole derivatives

The imidazole derivatives and their antimicrobial activities have been described previously [2,17]. Reported data on a series of 286 imidazole derivatives were used. The MIC values that represented the lowest concentration at which there was no visible growth of *S. pyogenes ATCC* 81 and *S. pyogenes ATCC* 101 were obtained.

Anti-*S. pyogenes* activity was measured in the tube dilution test, which is the standard method for determining levels of resistance to antimicrobial substances. Serial dilutions of the compounds were prepared in the Müller-Hinton medium obtaining the following concentrations: 500, 375, 187.5, 93.7, 46.9, 23.4, 11.7, 5.9, 2.9, 1.5, 0.75, 0.38, and 0.19 mg/l. After that, 0.1 mL of diluted 24-hour culture suspension was added to each of the tubes. Growth of *S. pyogenes* was determined visually after incubation for 24 hours at 37°C. The highest dilution without growth was considered as the MIC. Antifungal activity was measured as the logarithm of the reciprocal of minimum inhibitory concentration against *S. pyogenes*, log 1/MICexp.

2.2. Molecular modeling

The three-dimensional structures of the 286 imidazole derivatives at their cationic state were calculated using Gaussian 03, Revision D.01 (Gaussian, Wallingford CT, USA). All geometry optimizations were performed on isolated molecules applying semiempirical AM1 methods. These quantum mechanical calculations were also used to determine selected descriptors: HF–the heat of formation [kcal/mole], μ–dipole moment (debye), DELH–the energy difference [eV] between the frontier molecular orbitals, ELUMO – EHOMO, where ELUMO is the energy of the lowest unoccupied molecular orbital and EHOMO is the energy of the highest occupied molecular orbital. All these descriptors were calculated using the CACheWorkSystem Pro version 7.5.0.85 package (CACheWorkSystem Pro, Fujitsu, Oxford, Great Britain). Furthermore, the molecular descriptors were calculated using DRAGON for Windows version 5.5-2007 package (Talete, Milano, Italy). The Dragon descriptors included 22 different logical blocks. The total number of calculated descriptors was 3224.

Descriptor dimensionalities proposed by DRAGON were 0D, 1D, 2D, 3D and "others". The subset 0D referred to atom and bond type counts, 1D to fragment counts, 2D to topological and related descriptors, and 3D to all the descriptors that depended on the geometrical coordinates of the molecule atoms. The subset of "others" included charge descriptors and molecular properties.

Several descriptor groups were analysed. Constitutional descriptors included 0D-descriptors independent from molecular connectivity and conformations Mw – molecular weight, Mp – mean atomic polarizability (scaled on Carbon atom) and sum of atomic properties such as: Sv – sum of atomic van der Waals volumes (scaled on C atom), Se – sum of atomic Sanderson electronegatives (scaled on C atom), Sp – sum of atomic polarizabilities (scaled on C atom), Ss – sum of Kier-Hall electrotopological states. Topological descriptors (2D-descriptors) included molecular connectivity index (χ_0, χ_1, χ_2), average connectivity index (χ_{0A}, χ_{1A}, χ_{2A}), valence molecular connectivity index (χ_{0V}, χ_{1V}, χ_{2V})[a], index quantifying the shape of a chemical system (κ_1, κ_2, κ_3) [19,20]. These molecular descriptors were obtained from molecular graph, *i.e.,* 2D-descriptors conformationally independent. They were numerical quantifiers of molecular topology obtained by the application of algebraic operators to matrices representing molecular graphs of values independent of vertex numbering or labelling. They could be sensitive to one or more structural features of the molecule such as size, shape, symmetry, branching and cyclicity and could also encode chemical information concerning atom type and bond multiplicity. Weighted Holistic Invariant Molecular (WHIM) descriptors were geometrical descriptors (3D) obtained as statistical indices of the atoms projected onto the 3 principal components obtained from weighted covariance matrices of the atomic coordinates [21]. Calculating WHIM descriptors included volume total size index (unweighted or weighted by atomic: masses; Sanderson electronegatives; van der Waals volumes; polarizabilities; electrotopological states) and K global shape index (unweighted or weighted by atomic: polarizabilities; Sanderson electronegativities). Molecular properties (subset of "others" descriptors) were calculated from models together with some empirical descriptors that included *log P* (octanol-water partition coefficient), MR – molar refractivity, TPSA (Tot) – topological polar surface area (scaled on N, O, S, P) and TPSA (NO) – topological polar surface area (scald on N, O) [22,23].

Several criteria were used to reduce the number of descriptors while optimizing the information content of the descriptors set. First, descriptors for which no value was available for all the compounds were disregarded. Secondly, the descriptors showing the same value (or nearly the same) for all compounds were excluded. For the remaining descriptors, if two descriptors showed a correlation coefficient greater than 0.9, the one showing the highest pair correlation with the other descriptors was removed. Indeed, the threshold value of the correlation coefficient of 0.9 is somewhat high, but this is the lowest value included in the software to exclude correlated descriptors. Usually, when the number of descriptors obtained after the preliminary screening is high, further exclusion procedures are applied.

Finally, the chosen descriptors were used as regressors of the model. They are shown in Table 1, and their detailed description can be found in the literature [21-24]. The value of each descriptor for 286 imidazole derivatives used in ANN is listed in Table 1S.

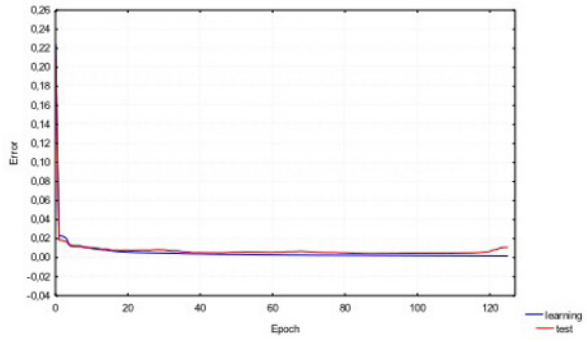
Table 1. Molecular descriptors used in ANN analysis.

No.	Variable	Block description
1	Heat of Formation (kcal/mole)	quantum-chemical descriptors
2	Dipole Moment (debye)	quantum-chemical descriptors
3	LUMO Energy (eV)	quantum-chemical descriptors
4	HOMO Energy (eV)	quantum-chemical descriptors
5	DELH	quantum-chemical descriptors
6	Connectivity Index (order 0, standard)	connectivity indices
7	Connectivity Index (order 1, standard)	connectivity indices
8	Connectivity Index (order 2, standard)	connectivity indices
9	Valence Connectivity Index (order 0, standard)	connectivity indices
10	Valence Connectivity Index (order 1, standard)	connectivity indices
11	Valence Connectivity Index (order 2, standard)	connectivity indices
12	Average connectivity index chi-0	connectivity indices
13	Average connectivity index chi-1	connectivity indices
14	Average connectivity index chi-2	connectivity indices
15	Log P - octanol-water partition. Coeff.	molecular properties
16	Molar refractivity	molecular properties
17	Moriguchi octanol-water partition coeff. (logP)	molecular properties
18	Ghose-Crippen molar refractivity	molecular properties
19	Topological polar surface area (scaled on N, O)	molecular properties
20	Topological polar surface area (scaled on N, O, S, P)	molecular properties
21	Shape Index (basic kappa, order 1)	topological descriptors
22	Shape Index (basic kappa, order 2)	topological descriptors
23	Shape Index (basic kappa, order 3)	topological descriptors
24	Molecular weight	constitutional descriptors
25	Sum of atomic van der Waals volumes (scaled on carbon atom)	constitutional descriptors
26	Sum of atomic Sanderson electronegatives (scaled on carbon atom)	constitutional descriptors
27	Sum of Kier-Hall electrotopological states	constitutional descriptors
28	Sum of atomic polarizabilities (scaled on carbon atom)	constitutional descriptors
29	Mean atomic polarizability (scaled on Carbon atom)	constitutional descriptors
30	Volume total size index/ unweighted	WHIM descriptors
31	Volume total size index/ weighted by atomic masse	WHIM descriptors
32	Volume total size index/ weighted by atomic van der Waals volumes	WHIM descriptors
33	Volume total size index/ weighted by atomic Sanderson electronegatives	WHIM descriptors
34	Volume total size index/ weighted by atomic polarizabilities	WHIM descriptors
35	Volume total size index/ weighted by atomic electrotopological states	WHIM descriptors
36	K global shape index / unweighted	WHIM descriptors
37	K global shape index / weighted by atomic polarizabilities	WHIM descriptors
38	K global shape index / weighted by atomic Sanderson electronegativities	WHIM descriptors

2.3. ANN analysis

We used a multilayer perceptron artificial neural network, which was performed with the use of Statistica v. 8.0 software (StatSoft, Inc., Tulsa, USA) with the Automated Artificial Neural Networks module (www.statsoft.com). To keep the network structure as simple as possible, the ANN consisted of three layers: an input layer with 38 neurons, a hidden layer adjusted experimentally that consisted of 8 artificial neurons of tanh (hyperbolic tangent) activation function and a single output neuron of exponential activation function. The total set of 286 imidazole analogs was randomly subdivided into three sets: training (202 objects), validation (42 objects) and testing (42 objects). Before this procedure, all data sets were scaled within 0-1 range. The training set was the

Figure 2. Learning curve of MLP: 38:8:1.

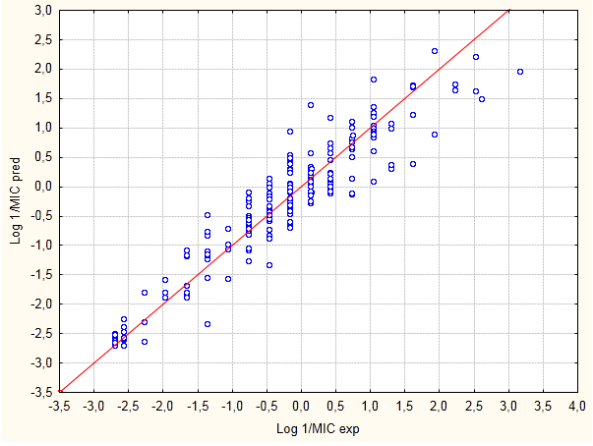


largest set of data, and it was subjected to the learning process. The algorithm that has been used for network training was Broyden-Fletcher-Goldfrab-Shanno BFGS (Quasi-Newton). After that, during a learning process, a test was performed to evaluate the generalization ability of a trained network. Next, a final check on the trained network was used with a validation set. The full training process was controlled by a root-mean-square (RMS) error reaching the smallest value with regard to the validation set of data. The program was stopped automatically when received the smallest value of RMS error. The most suitable network was found in epoch 87. The architecture and learning curves are presented in Figure 2. The validation set was used to ensure that there was no overfitting in the final results. The test set was designed to provide independent assessment of the network's performance when an entire procedure for network design was completed.

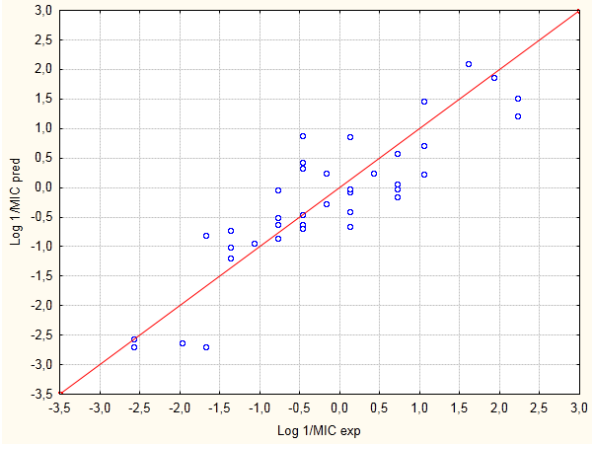
3. Results and discussion

The ANN model was projected, built up and trained. During the training process data from the learning sets were presented to the ANN to recognize connectivity between input and output signals. Correlation analysis is a technique use to measure the association between two variables. The results of correlation analysis obtained in this paper are presented in Figure 3. There are significant correlations between the theoretically calculated molecular descriptors and the experimentally determined antibacterial activity. The obtained correlation coeffi cient was 0.9461 for the learning set, 0.9060 for the validation set and 0.8824 for the testing set of imidazole derivatives. Additionally, the sensitivity analysis for input variables provided important information about the usefulness of individual variables [16,25]. On the basis of that analysis, the most significant predictive factors include WHIM, molecular indices and connectivity indices. The importance of usefulness of the individual descripors is listed in order of the rank in Table 2. The

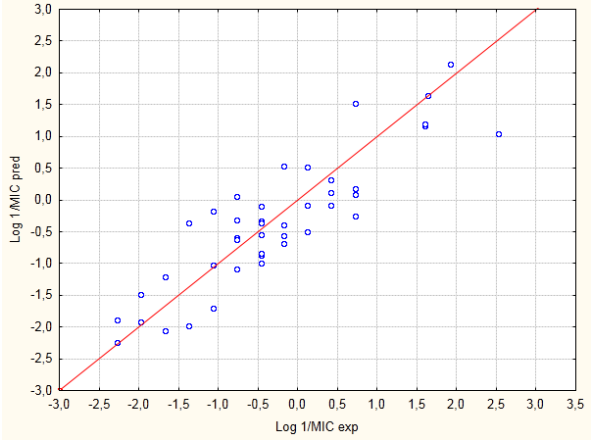
Figure 3. Correlation between the calculated and the experimental antifungal data for
a) learning R = 0.9461, RMSEP = 0.4047;
b) validation R = 0.9060,RMSEP = 0.5633;
c) testing set R = 0.8824, RMSEP = 0.5329.



a) learning R = 0.9461, RMSEP = 0.4047;



b) validation R = 0.9060,RMSEP = 0.5633;



c) testing set R = 0.8824, RMSEP = 0.5329;

Table 2. Sensitivity analysis.

Variables	Error	Rank
K global shape index / weighted by atomic polarizabilities	10.6972	1
K global shape index / unweighted	6.5916	2
K global shape index / weighted by atomic Sanderson electronegativities	4.6357	3
Topological polar surface area (scaled on N, O, S, P)	3.9053	4
Average connectivity index chi-0	3.1521	5
Moriguchi octanol-water partition coeff. (logP)	3.0478	6
Average connectivity index chi-1	2.4866	7
Shape Index (basic kappa, order 3)	2.3269	8
Topological polar surface area (scaled on N, O)	2.2705	9
Log P - octanol-water partition. coeff.	2.0316	10
Average connectivity index chi-2	1.7310	11
Mean atomic polarizability (scaled on Carbon atom)	1.6827	12
Heat of Formation (kcal/mole)	1.6669	13
Dipole Moment (debye)	1.6083	14
LUMO Energy (eV)	1.5362	15
HOMO Energy (eV)	1.5007	16
Valence Connectivity Index (order 2, standard)	1.3310	17
Sum of atomic Sanderson electronegatives (scaled on carbon atom)	1.2958	18
Connectivity Index (order 2, standard)	1.2854	19
Sum of atomic van der Waals volumes (scaled on carbon atom)	1.2623	20
Sum of atomic polarizabilities (scaled on carbon atom)	1.2507	21
DELH	1.2243	22
Connectivity Index (order 1, standard)	1.2176	23
Volume total size index/ weighted by atomic polarizabilities	1.2027	24
Volume total size index/ weighted by atomic masse	1.2017	25
Molecular weight	1.1999	26
Shape Index (basic kappa, order 2)	1.1845	27
Valence Connectivity Index (order 1, standard)	1.1523	28
Ghose-Crippen molar refractivity	1.1371	29
Connectivity Index (order 0, standard)	1.1276	30
Volume total size index/ weighted by atomic Sanderson electronegatives	1.1267	31
Molar Refractivity	1.1261	32
Volume total size index/ unweighted	1.1104	33
Valence Connectivity Index (order 0, standard)	1.1064	34
Sum of Kier-Hall electrotopological states	1.0969	35
Shape Index (basic kappa, order 1)	1.0965	36
Volume total size index/ weighted by atomic van der Waals volumes	1.0769	37
Volume total size index/ weighted by atomic electrotopological states	1.0489	38

most significant descriptors used with the rank close to 1 were K global shape index/weighted by atomic polarizabilities, K global shape index/unweighted, and K global shape index/weighted by atomic Sanderson electronegativities. The least significant with the rank close to 38 included volume total size index/weighted by atomic electrotopological states, volume total size index/weighted by atomic van der Waals volumes, and Shape Index (basic kappa, order 1). All the WHIM descriptors were built in such a way as to capture relevant molecular three-dimensional information regarding molecular shape, size, symmetry, and atom distribution with respect to the invariant reference frames. Their usefulness has been confirmed in such examples of modeling as toxicological indices, physiochemical properties of polichlorbiphenyls, polycyclic aromatic hydrocarbons,

hydroxyl radical rate constant and soil sorption partition coefficients [24]. The examples of modeling with the use of WHIM descriptors could suggest their further use in prediction of activities of other compounds, like the set with antimicrobial activities in our research.

Previous studies have shown the benefits of the use of ANN analysis in prediction of antimicrobial activity of different imidazole derivatives against many microorganisms [11,15,16]. Currently, the ANN predicted log1/MIC values were also significantly correlated with experimentally obtained log 1/MIC. The comparison of the predicted and experimental log 1/MIC as a means of difference between experimental and predicted log 1/MIC values (Δ) is presented in Table 1S.

4. Conclusion

The objective of our study was to demonstrate the possibility of constructing ANN for predictions of antimicrobial activity of imidazole derivatives against *Streptococcus pyogenes* of a defined structure. The results confirm the benefits of ANN analysis as a convenient tool to predict the log1/MIC for *S. pyogenes* with reference content of imidazole derivatives. The proposed method based on the database with a representative compounds group might be used for initial classification of biologically important elements and could be included in the set of methods of QSAR analysis.

Table 1S. Experimental and predicted logarithms of reciprocal of the minimal inhibitory concentrations, log 1/MIC along with the errors between those two values.

No.	Set	R1	R2	R3	R4	X	MIC[mg/dm^3]	log 1/MICexp	log 1/MICpred	Δ*
1	L	C4H9	-	C4H9	-	O	375.00	-2.5740	-2.5994	0.0253
2	L	C6H13	-	C4H9	-	O	46.90	-1.6712	-1.8814	0.2102
3	T	C8H17	-	C4H9	-	O	23.40	-1.3692	-0.3687	-1.0005
4	L	C10H21	-	C4H9	-	O	0.75	0.1249	0.3415	-0.2166
5	L	C12H25	-	C4H9	-	O	0.09	1.0458	0.0926	0.9532
6	T	C14H29	-	C4H9	-	O	11.70	-1.0682	-0.1737	-0.8945
7	L	C16H33	-	C4H9	-	O	11.70	-1.0682	-0.7089	-0.3593
8	T	C4H9	-	C6H13	-	O	46.90	-1.6712	-2.0580	0.3868
9	L	C6H13	-	C6H13	-	O	23.40	-1.3692	-1.5420	0.1728
10	L	C8H17	-	C6H13	-	O	1.50	-0.1761	0.0395	-0.2156
11	V	C10H21	-	C6H13	-	O	0.09	1.0458	0.7109	0.3348
12	L	C12H25	-	C6H13	-	O	0.19	0.7213	0.6503	0.0709
13	L	C14H29	-	C6H13	-	O	2.90	-0.4624	0.1447	-0.6071
14	T	C16H33	-	C6H13	-	O	5.90	-0.7709	-0.5920	-0.1788
15	L	C4H9	-	C8H17	-	O	23.40	-1.3692	-1.0902	-0.2790
16	V	C6H13	-	C8H17	-	O	0.75	0.1249	-0.0700	0.1950
17	L	C8H17	-	C8H17	-	O	0.09	1.0458	1.0482	-0.0024
18	L	C10H21	-	C8H17	-	O	0.03	1.6021	1.2303	0.3718
19	L	C12H25	-	C8H17	-	O	0.19	0.7213	0.8357	-0.1145
20	L	C14H29	-	C8H17	-	O	1.50	-0.1761	0.0666	-0.2427
21	L	C16H33	-	C8H17	-	O	11.70	-1.0682	-1.0604	-0.0078
22	L	C4H9	-	C10H21	-	O	0.38	0.4202	0.0736	0.3466
23	V	C6H13	-	C10H21	-	O	0.75	0.1249	0.8562	-0.7312
24	L	C8H17	-	C10H21	-	O	0.03	1.6021	1.7276	-0.1255
25	T	C10H21	-	C10H21	-	O	0.19	0.7213	1.5151	-0.7938
26	L	C12H25	-	C10H21	-	O	0.38	0.4202	0.1260	0.2943
27	T	C14H29	-	C10H21	-	O	5.90	-0.7709	-1.0933	0.3225
28	L	C16H33	-	C10H21	-	O	11.70	-1.0682	-1.5618	0.4936

Table 1S continued. Experimental and predicted logarithms of reciprocal of the minimal inhibitory concentrations, log 1/MIC along with the errors between those two values.

29	L	C4H9	-	C12H25	-	O	1.50	-0.1761	0.5527	-0.7288
30	T	C6H13	-	C12H25	-	O	0.03	1.6021	1.1667	0.4354
31	L	C8H17	-	C12H25	-	O	0.19	0.7213	0.7688	-0.0476
32	L	C10H21	-	C12H25	-	O	1.50	-0.1761	-0.2151	0.0390
33	L	C12H25	-	C12H25	-	O	1.50	-0.1761	-0.3548	0.1787
34	L	C14H29	-	C12H25	-	O	5.90	-0.7709	-0.6828	-0.0881
35	V	C16H33	-	C12H25	-	O	11.70	-1.0682	-0.9478	-0.1204
36	L	C4H9	CH3	CH3	H	O	500.00	-2.6990	-2.6983	-0.0007
37	L	C6H13	CH3	CH3	H	O	375.00	-2.5740	-2.4707	-0.1033
38	L	C8H17	CH3	CH3	H	O	46.90	-1.6712	-1.6697	-0.0015
39	T	C10H21	CH3	CH3	H	O	1.50	-0.1761	-0.6911	0.5150
40	L	C12H25	CH3	CH3	H	O	0.38	0.4202	-0.1080	0.5282
41	V	C4H9	C2H5	CH3	H	O	375.00	-2.5740	-2.6978	0.1238
42	L	C6H13	C2H5	CH3	H	O	46.90	-1.6712	-1.8058	0.1346
43	T	C8H17	C2H5	CH3	H	O	1.50	-0.1761	-0.5671	0.3911
44	V	C10H21	C2H5	CH3	H	O	0.75	0.1249	-0.0291	0.1540
45	L	C12H25	C2H5	CH3	H	O	0.38	0.4202	-0.0255	0.4457
46	V	C4H9	n-C3H7	CH3	H	O	375.00	-2.5740	-2.6978	0.1238
47	T	C6H13	n-C3H7	CH3	H	O	23.40	-1.3692	-1.9842	0.6150
48	L	C8H17	n-C3H7	CH3	H	O	5.90	-0.7709	-0.6070	-0.1639
49	T	C10H21	n-C3H7	CH3	H	O	0.38	0.4202	0.1232	0.2970
50	V	C12H25	n-C3H7	CH3	H	O	0.19	0.7213	-0.0187	0.7400
51	L	C4H9	iso-C3H7	CH3	H	O	500.00	-2.6990	-2.6977	-0.0013
52	L	C6H13	iso-C3H7	CH3	H	O	500.00	-2.6990	-2.4996	-0.1994
53	T	C8H17	iso-C3H7	CH3	H	O	93.70	-1.9717	-1.4907	-0.4811
54	L	C10H21	iso-C3H7	CH3	H	O	5.90	-0.7709	-0.7749	0.0040
55	L	C12H25	iso-C3H7	CH3	H	O	2.90	-0.4624	-0.4691	0.0067
56	L	C4H9	H	CH3	Cl	O	500.00	-2.6990	-2.6982	-0.0008
57	L	C6H13	H	CH3	Cl	O	500.00	-2.6990	-2.5767	-0.1223
58	L	C8H17	H	CH3	Cl	O	500.00	-2.6990	-2.5191	-0.1799
59	L	C10H21	H	CH3	Cl	O	375.00	-2.5740	-2.4679	-0.1062
60	L	C12H25	H	CH3	Cl	O	187.50	-2.2730	-2.3038	0.0308
61	V	C2H5	-	CH2OC2H5	-	O	375.00	-2.5740	-2.6982	0.1242
62	L	C4H9	-	CH2OC2H5	-	O	375.00	-2.5740	-2.6976	0.1236
63	L	C5H11	-	CH2OC2H5	-	O	375.00	-2.5740	-2.6961	0.1221
64	T	C6H13	-	CH2OC2H5	-	O	187.50	-2.2730	-1.8834	-0.3896
65	L	C7H15	-	CH2OC2H5	-	O	23.40	-1.3692	-1.2225	-0.1467
66	L	C8H17	-	CH2OC2H5	-	O	2.90	-0.4624	-0.5820	0.1196
67	V	C9H19	-	CH2OC2H5	-	O	1.50	-0.1761	-0.2688	0.0927
68	L	C10H21	-	CH2OC2H5	-	O	2.90	-0.4624	-0.1135	-0.3489
69	L	C12H25	-	CH2OC2H5	-	O	2.90	-0.4624	-0.2066	-0.2558
70	L	C14H29	-	CH2OC2H5	-	O	1.50	-0.1761	-0.6450	0.4689

Table 1S continued. Experimental and predicted logarithms of reciprocal of the minimal inhibitory concentrations, log 1/MIC along with the errors between those two values.

71	L	C16H33	-	CH2OC2H5	-	O	46.90	-1.6712	-1.1750	-0.4961
72	L	C4H9	CH2OC2H5	-	-	S	375.00	-2.5740	-2.2426	-0.3315
73	L	C6H13	CH2OC2H5	-	-	S	5.90	-0.7709	-0.8024	0.0315
74	L	C8H17	CH2OC2H5	-	-	S	0.75	0.1249	-0.2772	0.4021
75	L	C10H21	CH2OC2H5	-	-	S	1.50	-0.1761	-0.0051	-0.1710
76	L	C12H25	CH2OC2H5	-	-	S	2.90	-0.4624	0.0107	-0.4731
77	L	C2H5	-	CH2OC3H7	-	O	500.00	-2.6990	-2.6982	-0.0008
78	L	C4H9	-	CH2OC3H7	-	O	500.00	-2.6990	-2.6966	-0.0024
79	L	C6H13	-	CH2OC3H7	-	O	93.70	-1.9717	-1.7917	-0.1800
80	L	C8H17	-	CH2OC3H7	-	O	5.90	-0.7709	-0.4942	-0.2766
81	L	C10H21	-	CH2OC3H7	-	O	5.90	-0.7709	-0.0840	-0.6868
82	L	C12H25	-	CH2OC3H7	-	O	1.50	-0.1761	-0.2892	0.1131
83	L	C14H29	-	CH2OC3H7	-	O	1.50	-0.1761	-0.6942	0.5181
84	T	C16H33	-	CH2OC3H7	-	O	2.90	-0.4624	-0.9982	0.5358
85	L	C4H9	-	CH2OC3H7	-	S	93.70	-1.9717	-1.8800	-0.0917
86	L	C6H13	-	CH2OC3H7	-	S	2.90	-0.4624	-0.3784	-0.0840
87	L	C8H17	-	CH2OC3H7	-	S	0.19	0.7213	-0.1276	0.8488
88	L	C10H21	-	CH2OC3H7	-	S	0.75	0.1249	0.0357	0.0893
89	L	C12H25	-	CH2OC3H7	-	S	0.75	0.1249	-0.2335	0.3585
90	L	C3H9	-	CH2OC3H7	-	O	500.00	-2.6990	-2.6981	-0.0009
91	V	C5H11	-	CH2OC3H7	-	O	375.00	-2.5740	-2.5654	-0.0086
92	L	C7H15	-	CH2OC3H7	-	O	23.40	-1.3692	-1.1623	-0.2069
93	T	C9H19	-	CH2OC3H7	-	O	46.90	-1.6712	-1.2100	-0.4611
94	L	C4H9	C6H5	C8H17	-	S	0.38	0.4202	0.7396	-0.3194
95	L	C6H13	C6H5	C8H17	-	S	0.09	1.0458	1.3585	-0.3128
96	L	C8H17	C6H5	C8H17	-	S	0.01	2.2219	1.7486	0.4733
97	L	C10H21	C6H5	C8H17	-	S	0.38	0.4202	1.1772	-0.7570
98	L	C12H25	C6H5	C8H17	-	S	0.75	0.1249	0.2876	-0.1626
99	L	C16H33	C6H5	C8H17	-	S	0.75	0.1249	0.0192	0.1058
100	T	C3H7	C6H5	C8H17	-	O	0.38	0.4202	0.3105	0.1097
101	L	C5H11	C6H5	C8H17	-	O	0.09	1.0458	1.2600	-0.2143
102	L	C7H15	C6H5	C8H17	-	O	0.01	1.9208	2.3168	-0.3960
103	T	C9H19	C6H5	C8H17	-	O	0.00	2.5229	1.0453	1.4776
104	L	C11H21	C6H5	C8H17	-	O	0.05	1.3010	0.3153	0.9858
105	L	C2H5	C6H5	C8H17	-	O	0.19	0.7213	-0.1071	0.8284
106	L	C4H9	C6H5	C8H17	-	O	0.09	1.0458	0.9907	0.0551
107	L	C6H13	C6H5	C8H17	-	O	0.03	1.6021	1.6947	-0.0926
108	L	C8H17	C6H5	C8H17	-	O	0.01	2.2219	1.6551	0.5667
109	L	C10H21	C6H5	C8H17	-	O	0.01	1.9208	0.8935	1.0273
110	T	C12H25	C6H5	C8H17	-	O	0.19	0.7213	0.0801	0.6412
111	L	C14H29	C6H5	C8H17	-	O	0.38	0.4202	0.5714	-0.1512
112	L	C16H33	C6H5	C8H17	-	O	0.38	0.4202	0.0332	0.3870

Table 1S continued. Experimental and predicted logarithms of reciprocal of the minimal inhibitory concentrations, log 1/MIC along with the errors between those two values.

113	L	C5H11	C6H5	C8H17	-	S	0.19	0.7213	1.0075	-0.2862
114	L	CH2C6H5	C6H5	C8H17	-	S	0.19	0.7213	0.5098	0.2115
115	L	CH2C6H5	C6H5	C8H17	-	O	0.19	0.7213	1.1134	-0.3921
116	L	C2H5	C6H5	C6H13	-	O	2.90	-0.4624	-1.3298	0.8674
117	L	C4H9	C6H5	C6H13	-	O	2.90	-0.4624	-0.7233	0.2609
118	V	C6H13	C6H5	C6H13	-	O	2.90	-0.4624	0.3282	-0.7906
119	V	C8H17	C6H5	C6H13	-	O	2.90	-0.4624	0.8765	-1.3389
120	L	C10H21	C6H5	C6H13	-	O	0.00	2.5229	2.2103	0.3126
121	L	C12H25	C6H5	C6H13	-	O	0.03	1.6021	0.4001	1.2020
122	L	C14H29	C6H5	C6H13	-	O	1.50	-0.1761	-0.0844	-0.0917
123	T	C16H33	C6H5	C6H13	-	O	2.90	-0.4624	-0.3294	-0.1330
124	L	C4H9	C6H5	C6H13	-	S	1.50	-0.1761	0.0189	-0.1950
125	L	C6H13	C6H5	C6H13	-	S	0.19	0.7213	0.6811	0.0401
126	L	C8H17	C6H5	C6H13	-	S	0.00	2.5229	1.6239	0.8990
127	L	C10H21	C6H5	C6H13	-	S	0.00	3.1549	1.9694	1.1855
128	L	C12H25	C6H5	C6H13	-	S	0.00	2.6021	1.5006	1.1015
129	T	C14H29	C6H5	C6H13	-	S	0.19	0.7213	0.1765	0.5447
130	L	C16H33	C6H5	C6H13	-	S	0.75	0.1249	0.0508	0.0742
131	L	C3H7	C6H5	C6H13	-	O	5.90	-0.7709	-1.2635	0.4927
132	L	C5H11	C6H5	C6H13	-	O	2.90	-0.4624	-0.5228	0.0604
133	V	C7H15	C6H5	C6H13	-	O	0.75	0.1249	-0.4061	0.5310
134	L	C9H19	C6H5	C6H13	-	O	1.50	-0.1761	0.9538	-1.1299
135	L	C11H23	C6H5	C6H13	-	O	1.50	-0.1761	0.4438	-0.6199
136	T	iso-C4H9	C6H5	C6H13	-	O	11.70	-1.0682	-1.6983	0.6301
137	L	CH2C6H5	C6H5	C6H13	-	O	2.90	-0.4624	-0.1576	-0.3048
138	L	C5H11	C6H5	C6H13	-	O	0.75	0.1249	0.5776	-0.4526
139	L	C2H5	C6H5	H	-	O	375.00	-2.5740	-2.5802	0.0062
140	T	C4H9	C6H5	H	-	O	187.50	-2.2730	-2.2443	-0.0287
141	L	C6H13	C6H5	H	-	O	23.40	-1.3692	-1.0846	-0.2846
142	L	C8H17	C6H5	H	-	O	1.50	-0.1761	-0.4575	0.2814
143	T	C10H21	C6H5	H	-	O	1.50	-0.1761	-0.3920	0.2159
144	L	C12H25	C6H5	H	-	O	23.40	-1.3692	-0.4707	-0.8986
145	L	C14H29	C6H5	H	-	O	5.90	-0.7709	-1.0819	0.3111
146	L	C16H33	C6H5	H	-	O	46.90	-1.6712	-1.0733	-0.5979
147	V	C2H5	C6H5	C4H9	-	O	46.90	-1.6712	-2.6924	1.0212
148	L	C4H9	C6H5	C4H9	-	O	23.40	-1.3692	-1.0997	-0.2695
149	L	C6H13	C6H5	C4H9	-	O	23.40	-1.3692	-0.8233	-0.5459
150	T	C8H17	C6H5	C4H9	-	O	5.90	-0.7709	0.0536	-0.8244
151	L	C10H21	C6H5	C4H9	-	O	1.50	-0.1761	0.0140	-0.1900
152	L	C12H25	C6H5	C4H9	-	O	0.75	0.1249	0.0044	0.1205
153	T	C14H29	C6H5	C4H9	-	O	0.75	0.1249	-0.5067	0.6317
154	L	C16H33	C6H5	C4H9	-	O	1.50	-0.1761	-0.5906	0.4145

Table 1S continued. Experimental and predicted logarithms of reciprocal of the minimal inhibitory concentrations, log 1/MIC along with the errors between those two values.

155	L	C3H7	C6H5	H	-	O	187.50	-2.2730	-2.6288	0.3558
156	L	C5H11	C6H5	H	-	O	46.90	-1.6712	-1.8723	0.2012
157	L	C7H15	C6H5	H	-	O	5.90	-0.7709	-1.0423	0.2715
158	L	C9H19	C6H5	H	-	O	2.90	-0.4624	-0.8168	0.3544
159	V	C3H7	C6H5	C4H9	-	O	93.70	-1.9717	-2.6262	0.6545
160	V	C5H11	C6H5	C4H9	-	O	23.40	-1.3692	-1.1927	-0.1766
161	T	C7H15	C6H5	C4H9	-	O	5.90	-0.7709	-0.3162	-0.4546
162	L	C9H19	C6H5	C4H9	-	O	1.50	-0.1761	0.3043	-0.4803
163	L	C4H9	C6H5	H	-	S	2.90	-0.4624	-0.3298	-0.1326
164	V	C6H13	C6H5	H	-	S	0.09	1.0458	1.4633	-0.4175
165	L	C8H17	C6H5	H	-	S	0.09	1.0458	0.9395	0.1062
166	L	C10H21	C6H5	H	-	S	0.75	0.1249	-0.1254	0.2503
167	L	C12H25	C6H5	H	-	S	0.38	0.4202	-0.0659	0.4861
168	L	C4H9	C6H5	C4H9	-	S	1.50	-0.1761	-0.2593	0.0832
169	L	C6H13	C6H5	C4H9	-	S	0.05	1.3010	0.3788	0.9222
170	L	C8H17	C6H5	C4H9	-	S	0.09	1.0458	0.9197	0.1261
171	V	C10H21	C6H5	C4H9	-	S	0.01	2.2219	1.5069	0.7150
172	L	C12H25	C6H5	C4H9	-	S	0.09	1.0458	1.8370	-0.7912
173	V	C4H9	C6H5	C10H21	-	S	0.03	1.6021	2.0933	-0.4912
174	T	C5H11	C6H5	C10H21	-	S	0.01	1.9208	2.1381	-0.2173
175	V	C6H13	C6H5	C10H21	-	S	0.01	1.9208	1.8675	0.0533
176	L	C8H17	C6H5	C10H21	-	S	0.09	1.0458	0.8771	0.1687
177	L	C10H21	C6H5	C10H21	-	S	0.38	0.4202	0.4669	-0.0466
178	L	C12H25	C6H5	C10H21	-	S	1.50	-0.1761	0.0567	-0.2328
179	L	C14H29	C6H5	C10H21	-	S	1.50	-0.1761	-0.0131	-0.1630
180	L	C16H33	C6H5	C10H21	-	S	2.90	-0.4624	0.0540	-0.5164
181	L	CH2C6H5	C6H5	C10H21	-	S	0.03	1.6021	1.7163	-0.1142
182	L	CH2C6H5	C6H5	C10H21	-	O	0.05	1.3010	1.0041	0.2970
183	L	C2H5	C6H5	C10H21	-	O	0.38	0.4202	0.0682	0.3520
184	L	C3H7	C6H5	C10H21	-	O	0.38	0.4202	0.0139	0.4063
185	L	C4H9	C6H5	C10H21	-	O	0.19	0.7213	0.1441	0.5772
186	L	C5H11	C6H5	C10H21	-	O	0.09	1.0458	0.8423	0.2034
187	V	C6H13	C6H5	C10H21	-	O	0.01	2.2219	1.2172	1.0047
188	T	C7H15	C6H5	C10H21	-	O	0.03	1.6021	1.1894	0.4126
189	L	C8H17	C6H5	C10H21	-	O	0.09	1.0458	1.1932	-0.1475
190	L	C9H19	C6H5	C10H21	-	O	0.75	0.1249	0.1776	-0.0526
191	L	C10H21	C6H5	C10H21	-	O	1.50	-0.1761	0.4252	-0.6013
192	V	C11H23	C6H5	C10H21	-	O	2.90	-0.4624	0.4252	-0.8876
193	L	C12H25	C6H5	C10H21	-	O	2.90	-0.4624	-0.5371	0.0747
194	T	C14H29	C6H5	C10H21	-	O	0.19	0.7213	-0.2584	0.9797
195	L	C16H33	C6H5	C10H21	-	O	2.90	-0.4624	-0.3225	-0.1399
196	L	C2H5	C6H5	C2H5	-	O	500.00	-2.6990	-2.6959	-0.0031

Table 1S continued. Experimental and predicted logarithms of reciprocal of the minimal inhibitory concentrations, log 1/MIC along with the errors between those two values.

197	L	C4H9	C6H5	C2H5	-	O	500.00	-2.6990	-2.6198	-0.0792
198	L	C6H13	C6H5	C2H5	-	O	46.90	-1.6712	-1.1640	-0.5071
199	L	C8H17	C6H5	C2H5	-	O	2.90	-0.4624	-0.4274	-0.0350
200	L	C10H21	C6H5	C2H5	-	O	5.90	-0.7709	-0.2478	-0.5231
201	L	C12H25	C6H5	C2H5	-	O	2.90	-0.4624	-0.4716	0.0092
202	L	C14H29	C6H5	C2H5	-	O	5.90	-0.7709	-0.7463	-0.0245
203	L	C16H33	C6H5	C2H5	-	O	23.40	-1.3692	-0.7503	-0.6190
204	L	C3H7	C6H5	C2H5	-	O	375.00	-2.5740	-2.6945	0.1205
205	L	C5H11	C6H5	C2H5	-	O	187.50	-2.2730	-1.7918	-0.4812
206	V	C7H15	C6H5	C2H5	-	O	23.40	-1.3692	-0.7332	-0.6360
207	L	C9H19	C6H5	C2H5	-	O	5.90	-0.7709	-0.0869	-0.6840
208	L	C11H23	C6H5	C2H5	-	O	5.90	-0.7709	-0.1913	-0.5796
209	L	C4H9	C6H5	C2H5	-	S	46.90	-1.6712	-1.7894	0.1183
210	L	C6H13	C6H5	C2H5	-	S	0.73	0.1367	-0.0893	0.2260
211	L	C8H17	C6H5	C2H5	-	S	0.73	0.1367	0.3174	-0.1808
212	L	C10H21	C6H5	C2H5	-	S	0.18	0.7447	0.8821	-0.1374
213	T	C12H25	C6H5	C2H5	-	S	0.02	1.6383	1.6417	-0.0034
214	L	C4H9	C6H5	C12H25	-	S	0.75	0.1249	1.3908	-1.2659
215	L	C5H11	C6H5	C12H25	-	S	0.38	0.4202	0.6614	-0.2411
216	V	C6H13	C6H5	C12H25	-	S	0.19	0.7213	-0.1603	0.8815
217	V	C8H17	C6H5	C12H25	-	S	0.19	0.7213	0.5750	0.1462
218	L	C10H21	C6H5	C12H25	-	S	0.75	0.1249	0.2486	-0.1237
219	L	C12H25	C6H5	C12H25	-	S	0.75	0.1249	0.2506	-0.1256
220	L	C14H29	C6H5	C12H25	-	S	0.75	0.1249	0.1216	0.0033
221	L	C16H33	C6H5	C12H25	-	S	1.50	-0.1761	0.0516	-0.2277
222	V	C2H5	C6H5	C12H25	-	O	5.90	-0.7709	-0.0421	-0.7287
223	L	C4H9	C6H5	C12H25	-	O	1.50	-0.1761	0.3876	-0.5637
224	T	C6H13	C6H5	C12H25	-	O	0.75	0.1249	0.5111	-0.3861
225	L	C8H17	C6H5	C12H25	-	O	1.50	-0.1761	0.4918	-0.6679
226	V	C10H21	C6H5	C12H25	-	O	1.50	-0.1761	0.2408	-0.4169
227	L	C12H25	C6H5	C12H25	-	O	2.90	-0.4624	-0.2122	-0.2502
228	T	C14H29	C6H5	C12H25	-	O	2.90	-0.4624	-0.1050	-0.3574
229	T	C16H33	C6H5	C12H25	-	O	2.90	-0.4624	-0.8739	0.4115
230	L	C3H7	C6H5	C12H25	-	O	1.50	-0.1761	-0.1851	0.0090
231	L	C5H11	C6H5	C12H25	-	O	0.75	0.1249	0.0866	0.0384
232	L	C7H15	C6H5	C12H25	-	O	1.50	-0.1761	0.2972	-0.4733
233	L	C9H19	C6H5	C12H25	-	O	2.90	-0.4624	-0.0045	-0.4579
234	L	C11H23	C6H5	C12H25	-	O	1.50	-0.1761	-0.0043	-0.1718
235	L	CH2C6H5	C6H5	C12H25	-	S	0.09	1.0458	0.6069	0.4389
236	L	CH2C6H5	C6H5	C12H25	-	O	0.09	1.0458	0.8983	0.1475
237	L	C2H5	C4H9	-	-	S	375.00	-2.5740	-2.5651	-0.0089
238	V	C2H5	C6H13	-	-	S	23.40	-1.3692	-1.0099	-0.3593

Table 1S continued. Experimental and predicted logarithms of reciprocal of the minimal inhibitory concentrations, log 1/MIC along with the errors between those two values.

239	L	C2H5	C8H17	-	-	S	5.90	-0.7709	-0.5281	-0.2428
240	L	C2H5	C10H21	-	-	S	1.50	-0.1761	-0.0070	-0.1691
241	L	C2H5	C12H25	-	-	S	0.05	1.3010	1.0745	0.2265
242	T	C2H5	C16H33	-	-	S	1.50	-0.1761	0.5305	-0.7066
243	L	C2H5	C2H5	-	-	O	375.00	-2.5740	-2.6972	0.1232
244	L	C2H5	C4H9	-	-	O	500.00	-2.6990	-2.6526	-0.0463
245	T	C2H5	C6H13	-	-	O	93.70	-1.9717	-1.9175	-0.0542
246	V	C2H5	C8H17	-	-	O	46.90	-1.6712	-0.8060	-0.8651
247	L	C2H5	C10H21	-	-	O	5.90	-0.7709	-0.3168	-0.4540
248	V	C2H5	C12H25	-	-	O	0.19	0.7213	0.0609	0.6604
249	L	C2H5	C14H29	-	-	O	0.38	0.4202	0.1191	0.3011
250	V	C2H5	C16H33	-	-	O	0.09	1.0458	0.2218	0.8240
251	L	n-C4H9	C4H9	-	-	S	93.70	-1.9717	-1.5702	-0.4015
252	V	n-C4H9	C6H13	-	-	S	5.90	-0.7709	-0.6251	-0.1457
253	L	n-C4H9	C8H17	-	-	S	1.50	-0.1761	-0.3124	0.1363
254	L	n-C4H9	C10H21	-	-	S	0.38	0.4202	0.2444	0.1758
255	L	n-C4H9	C12H25	-	-	S	0.38	0.4202	0.2327	0.1876
256	T	n-C4H9	C14H29	-	-	S	2.90	-0.4624	-0.3670	-0.0954
257	L	n-C4H9	C16H33	-	-	S	0.75	0.1249	-0.1220	0.2470
258	V	n-C4H9	C2H5	-	-	O	375.00	-2.5740	-2.6951	0.1211
259	L	n-C4H9	C4H9	-	-	O	375.00	-2.5740	-2.3860	-0.1880
260	T	n-C4H9	C6H13	-	-	O	11.70	-1.0682	-1.0317	-0.0365
261	T	n-C4H9	C8H17	-	-	O	5.90	-0.7709	-0.6206	-0.1502
262	T	n-C4H9	C10H21	-	-	O	0.38	0.4202	-0.0766	0.4968
263	L	n-C4H9	C12H25	-	-	O	0.38	0.4202	0.1497	0.2705
264	L	n-C4H9	C14H29	-	-	O	1.50	-0.1761	0.2423	-0.4184
265	V	n-C4H9	C16H33	-	-	O	1.50	-0.1761	-0.2784	0.1023
266	L	C6H13	C4H9	-	-	S	11.70	-1.0682	-0.9763	-0.0919
267	V	C6H13	C6H13	-	-	S	2.90	-0.4624	-0.4496	-0.0128
268	L	C6H13	C8H17	-	-	S	0.75	0.1249	-0.0776	0.2026
269	L	C6H13	C10H21	-	-	S	1.50	-0.1761	-0.4059	0.2298
270	L	C6H13	C12H25	-	-	S	5.90	-0.7709	-0.5607	-0.2102
271	L	C6H13	C2H5	-	-	O	23.40	-1.3692	-2.3233	0.9541
272	L	C6H13	C4H9	-	-	O	23.40	-1.3692	-1.1355	-0.2337
273	L	C6H13	C6H13	-	-	O	2.90	-0.4624	-0.8732	0.4108
274	V	C8H17	C4H9	-	-	S	0.75	0.1249	-0.6648	0.7897
275	L	C8H17	C6H13	-	-	S	1.50	-0.1761	-0.1900	0.0139
276	L	C8H17	C8H17	-	-	S	5.90	-0.7709	-0.7145	-0.0564
277	V	C8H17	C10H21	-	-	S	5.90	-0.7709	-0.8651	0.0943
278	V	C8H17	C12H21	-	-	S	2.90	-0.4624	-0.6320	0.1696
279	T	C8H17	C2H5	-	-	O	2.90	-0.4624	-0.8457	0.3833
280	V	C8H17	C4H9	-	-	O	2.90	-0.4624	-0.6847	0.2223

Table 1S continued. Experimental and predicted logarithms of reciprocal of the minimal inhibitory concentrations, log 1/MIC along with the errors between those two values.

281	L	C8H17	C6H13	-	-	O	1.50	-0.1761	-0.0744	-0.1017
282	T	C8H17	C8H17	-	-	O	0.75	0.1249	-0.0873	0.2122
283	V	C8H17	C10H21	-	-	O	0.38	0.4202	0.2412	0.1790
284	L	C8H17	C12H25	-	-	O	0.75	0.1249	-0.1457	0.2706
285	T	C8H17	C14H29	-	-	O	2.90	-0.4624	-0.5443	0.0819
286	V	C8H17	C16H33	-	-	O	5.90	-0.7709	-0.5141	-0.2568

(1- 236 imidazole rings Fig. 1a; 237-286 benzimidazole rings Fig.2b)
(Str 7 – Difference between experimental and predicted log 1/MIC values.
L – Learing set, V – Validation set, T – Testing set

References

[1] Weglewski J, Pernak J, Krysiński J,Synthesis and bactericidalproperties of pyridiniumchlorides with alkylthiomethyl and alkoxymethylhydrophobic-groups, J. Pharm.Sci.,1991,80, 91-95

[2] Pernak J, Krysiński J, Skrzypczak A,Activity of New Quaternary Iminium Compounds : [Bakterizide Wirkung von Iminiumverbindungen],Tenside Surfactants Detergents,1987, 24, 276-279

[3] Wiktorowicz W, Markuszewski M, Krysiński J, Kaliszan R, Quantitative structure-activity relationships study of a series of imidazole derivatives as potential new antifungal drugs,Acta Pol.Pharm., 2002, 59, 295-306

[4] Ranke J, Möltera K, Stocka F, Bottin-Webera U, Poczobutta J, Hoffmannb J, OndruschkabB, Filsera J, Jastorffa B, Biological effects of imidazolium ionic liquids with varying chain lengths in acute Vibrio fischeri and WST-1 cell viability assays,Ecotox. Env.Safety,2004,58, 396-404

[5] Hadj-esfandiari N, Navidpour L, Shadnia H, Amini M, Samadi N, Faramarzi MA, Shafiee A, Synthesis, antibacterial activity, and quantitative structure–activity relationships of new (Z)-2-(nitroimidazolylmethylene)-3-benzofuranone derivatives,Bioorg.Med. Chem.Lett.,2007,17, 6354-6363

[6] Ghasemi A, Namaki S, Mirshafiey A, Streptococcus pyogenes. J. Chin. Clin.Medicine,2009,4, 9-18

[7] Guilherme L, Postol E, Freschi de Barros S, Higa F, Alencar R, Lastre M, Zayas C, Puschel CR, Silva WR, Sa-Rocha LC, Sa-Rocha VM, Pérez O, Kalil J, A vaccine against S. pyogenes: Design and experimental immune response, Methods,2009, 49, 316-321

[8] Moura H, Woolfitt AR, Carvalho MG, Pavlopoulos A, Teixeira LM, Satten G A, Barr JR, MALDI-TOF mass spectrometryas a tool for differentiation of invasive and noninvasive Streptococcus pyogenes isolates, FEMS Immunol.Med.Microbiol.,2008,53, 333-342

[9] JR, MALDI-TOF mass spectrometryas a tool for differentiation of invasive and noninvasive Streptococcus pyogenes isolates, FEMS Immunol. Med.Microbiol.,2008,53, 333-342

[10] Miller AK, Antibacterial activity of ronidazole,Appl. Microbiol., 1971, 22, 480-481

[11] BerkelhammerG, Asato G, 2-Amino-5-(1-methyl-5-nitro-2-imidazolyl)-1,3, 4- thiadiazole: a new antimicrobial agent,Science,1986, 162, 1146

[12] Buciński A, Markuszewski MJ, Wiktorowicz W, Krysinski J, Kaliszan R, Artificial neural networks for prediction of antibacterial activity in series of imidazole derivatives, Comb. Chem. High ThroughputScreen,2004,7, 327-336

[13] Buciński A, Nasal A, Kaliszan R, Pharmacological classification of drugs based on neural network processing of molecular modeling data, Comb. Chem. High ThroughputScreen, 2000,3, 525-533

[14] Nasal A, Buciński A, Bączek T, Wojdełko A, Prediction of the affinity of the newly synthesisedazapirone derivatives for 5-HT1A receptors based on artificial neural network analysis of chromatographic retention data and calculation chemistry parameters, Comb.Chem. High ThroughputScreen,2004, 7, 313-325.

[15] So SS, Richards WG, Application of neural networks: quantitative structure-activity relationships of the derivatives of 2,4-diamino-5-(substituted-benzyl)pyrimidines as DHFR inhibitors,J.Med. Chem.,1992,35, 3201-3207

[16] Jaen-Oltra J, Salabert-Salvador MT, Garcia-March FJ, Perez-Gimenez F, Tomas-Vert F, Artificial neural network applied to prediction of fluorquinolone antibacterial activity by topological methods,

J.Med.Chem.,2000,43, 1143-1148
[17] Buciński A, Socha A, Wnuk M, Bączek T, Nowaczyk A, Krysiński J, Goryński K, Koba M, Artificial neural networks in prediction of antifungal activity of series of pyridine derivatives against Candida albicans, J. Microbiol. Methods,2009,79, 25-29
[18] Kövesdi I, Dominguez-Rodriguez MF, Ôrfi L, Náray-Szabó G, Varró A, Papp JG, Mátyus P, Application of neural networks in structure – activity relationships,Med. Res.Rev.,1999,19, 249-269
[19] Pernak J, Jędraszczyk J, Skrzypczak A, Krysiński J, Węcłaś H,Antimikrobiellewirkungvon iminiumverbindungen,MittTenside Surf. Det., 1992, 29, 280-284
[20] Kier LB, Hall LH, Molecular Connectivity in Structure-Activity Analysis, John Wiley and Sons, New York, 1986
[21] Hall LH, Kier LB, Reviews in Computational Chemistry, Ch. 9, ed. KB,Lipkowitz and DB,Boyd, 1992
[22] Todeschini R, Lasagni M, Marengo E, New Molecular Descriptors for 2D- and 3D-Structures, Theoretical J.Chemometr.,1984, 8, 263-273
[23] Moriguchi I, Hirono S, Liu Q, Nakagome Y, Matsushita Y, Simple method of calculating octanol/water partition coefficient, Chem. Pharm. Bull.,1992,40, 127- 130
[24] Ghose AK, Pritchett A, Crippen GM, Atomic physicochemical parameters for three dimensional structure directed quantitative structure-activity relationships III: Modeling hydrophobic interactions,J. Comp. Chem., 1988, 9, 80-90
[25] TodeschiniR, Consonni V, Handbook of molecular descriptors, Wiley-VHC, Weinheim (Germany), 2000
[26] Bączek T, Buciński A, Ivanow AR, Kaliszan R, Artificial Neural Network analysis for evaluation of peptide MS/MS spectra in proteomics, Anal. Chem.,2004,76, 1726-173